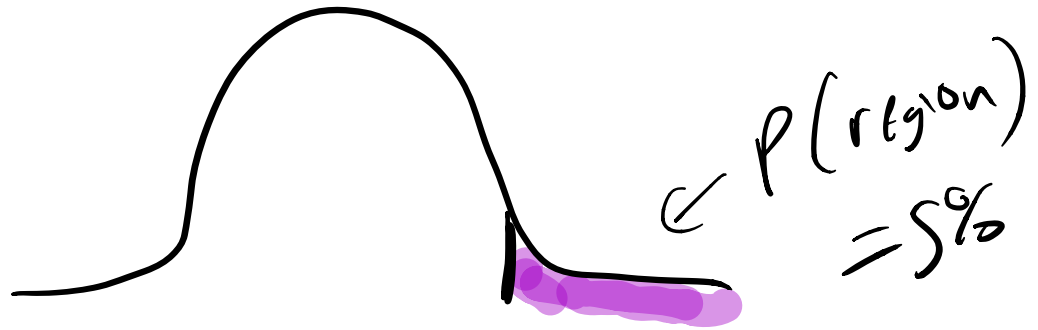# Nuance about p-values

P-value definition

The probability under the null hypothesis of seeing an as extreme or more extreme result.

If the null hypothesis is true then we falsely reject 5% of the time (if our significance level is 5%).

$\leftarrow P(\text{region})$

$= 5\%$

	null true	null false
fail to reject	✓	bad
reject	bad	✓

worst one

"sending an innocent person to jail"

Even if the null
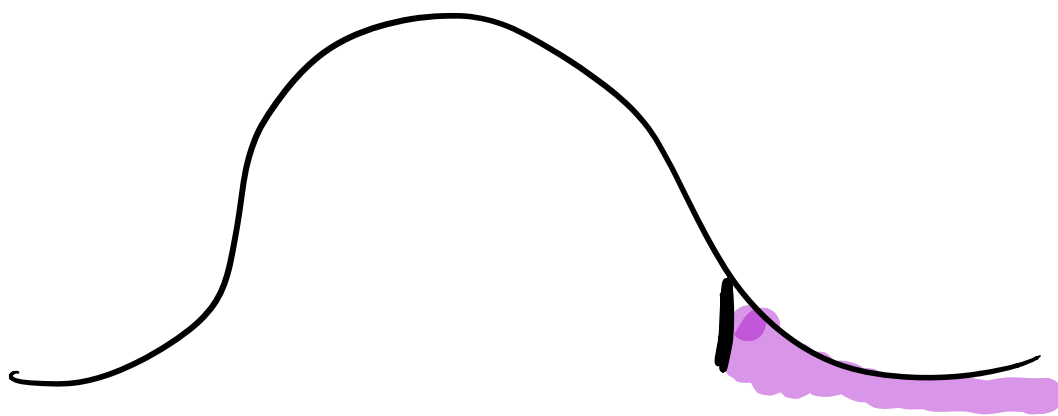is true we have
5% chance of falsely
rejecting.

Problem of multiple
testing, as we do
more tests probability
of false rejection
increases even if null is true.
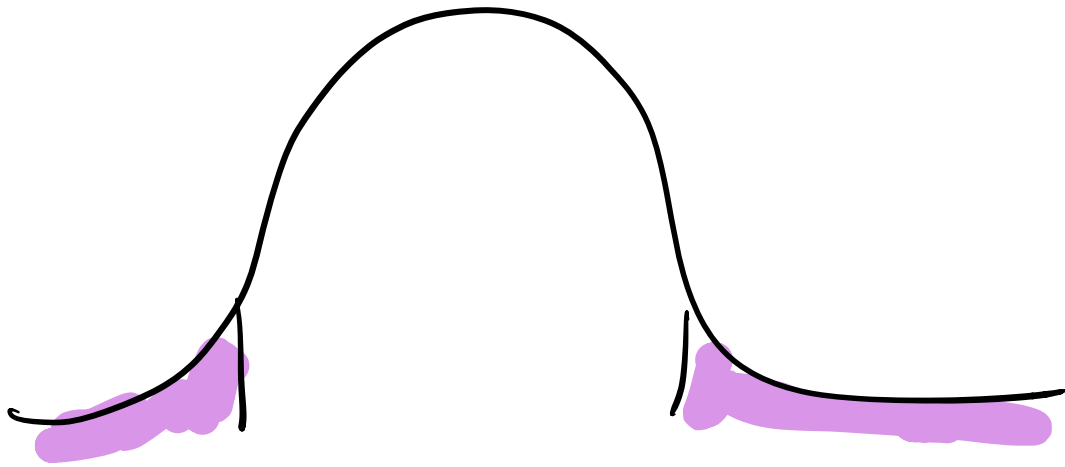
Two-Sided testing

Before:

H_0: $\alpha = 0$.

H_1: $\alpha > 0$.

H_0: $a = 0$.

H_1: $a \neq 0$.

p-value is doubled.

Q1 worksheet

	M	sep/Div/w	NM	
E	790	98	209	1097
UN	56	11	27	94
NL	21	7	13	41
	867	116	249	1232

H_0: Employment status is ind. of marital status

H_i: dependence.

$$P(M|E) = P(M).$$

$$= \frac{867}{1232}$$

EV table

	M		SeP		NM	
E	$\frac{867}{1232} \cdot 1097$		$\frac{116}{1232} \cdot 1097$		$\frac{289}{1232} \cdot 1097$	
UN	11	.94	11	94	11	.94
NC	11	.41	11	41	11	.41

$$\sum_{\substack{\text{all} \\ \text{entries}}} \frac{(\text{obs} - \text{exp})^2}{\exp}$$

$$df = (m-1)(n-1)$$

$$= 2 \cdot 2 = 4$$